



ДААЛГАВРЫН ХАРИУЛТЫН ОНОЛ (ITEM RESPONSE THEORY)-ЫН 3 ПАРАМЕТРТЭЙ ЗАГВАРЫГ АШИГЛАЖ ТЕСТИЙН ДААЛГАВРУУДАД ШИНЖИЛГЭЭ ХИЙХ НЬ

Х.Отгонбаатар¹

¹ Боловсролын хүрээлэн, Монгол улс
Цахим шуудан: otgonbaatar@mier.mn

Хураангуй

Энэхүү өгүүлэл нь Даалгаврын хариултын онол (Item Response Theory)-ын хэрэглээний ач холбогдол, тестийн даалгавар боловсруулах аргачлалыг тайлбарлахыг зорив. Ингэхдээ 7 дугаар ангийн сурагчдад зориулсан 4 сонголт бүхий математикийн тест боловсруулж, 480 сурагчаар 40 минутад ажиллуулж хариултуудыг цуглуулсан болно. Тестийн даалгавруудад өгсөн сурагчдын хариултуудад 3 параметртэй (a , b , c) IRT загвар, BILOG-MG программыг ашиглаж анализ хийв. Даалгаврын параметруудийг даалгаврын онцлогийн муруйг (Item Characteristic Curve) ашиглан тайлбарласан болно.

Түлхүүр үг: даалгаврын параметр, даалгаврын дискреминаци, даалгаврын онцлогийн муруй, IRT

ОРШИЛ

Америкийн эрдэмтэн Л.Л.Төрстоун 1925 онд хэвлүүлсэн “Сэтгэл судлал, сургалтын тестийг хэмжээсжүүлэх арга зүй” сэдэвт өгүүлэлдээ Орчин үеийн психометрийн онол болох Даалгаврын хариултын онол (Item Response Theory)-ын ухагдахуун, суурь ойлголтыг анхлан гаргаж тавьжээ. Хэдийгээр Л.Л.Төрстоун энэ чиглэлийн судалгаанаасаа бага зэрэг хазайж өөр чиглэлд судалгаа хийх болсон ч түүний шавь, хамтран ажиллагсад эл онолын үзэл баримтлал, аргачлалыг боловсронгуй болгох асуудалд ихээхэн хүчин чармайлт гаргасны дотор Ричардсон (1936), Фергуссон (1943), Фред Лорд (1952), Жорж Раш нарын эрдэмтэд IRT-ийн өөр өөр төрлийн моделиудыг боловсруулж, орчин үеийн судалгааны томоохон онол, аргачлал болгоход үнэтэй хувь нэмэр оруулсан юм. Өнөө үед боловсролын чанарын үнэлгээг

тодорхойлдог PISA, TIMSS, NAEP олон улсын сорилууд энэ онолд үндэслэж байна. Даалгаврын хариултын онол нь тест боловсруулах, тестийг үнэлэх хамгийн үр дүнтэй арга бөгөөд шалгуулагчийн чадвар, даалгаврын сонголт, тестийн загвар зэргийг үнэлэхэд маш чухал хэрэгсэл болж чаддаг юм. Энэ онол нь шалгуулагчийн чадварыг тооцож гаргахаас гадна тестийн даалгаврууд чухам хэр сайн болсон байна гэдгийг үнэлэх боломж олгодог давуу талтай юм. Өөрөөр хэлбэл тухайн даалгаврын онцлогийг илтгэдэг муруй (Item Characteristic Curve)-г ашиглаж шалгуулагчийн чадвар ба даалгаврын гүйцэтгэл хоорондын хамаарлыг харуулдаг. Шалгуулагчийн чадвар, даалгаврын параметрийг шалгуулагчийн тестэд үзүүлсэн хариултууд дээр үндэслэж тооцдог бөгөөд даалгаврын параметруудийн тоо нь

ямар загвар ашиглаж байгаагаас шалтгаалан өөр өөр байдаг. Хэдийгээр эдгээр загварууд нь математик илэрхийллүүдийг агуулсан байдаг ч суурь ойлголтуудыг нь ойлгоход тийм ч хэцүү биш. Хамгийн нийтлэг хэрэглэгддэг загварууд 1,2,3 (1PL, 2PL, 3PL) параметртэй загварууд юм. Энэхүү судалгааны ажилд доор тайлбарласан 3 параметртэй (3PL) загварыг ашигласан болно.

3 параметртэй IRT загвар (3PL)

1980 онд Германы эрдэмтэн Фред Лорд энэхүү 3 параметрт загварыг анх сонголттой хариулттай тестийн даалгаварт тааж хариулах боломжийг тооцож боловсруулсан байна. Даалгавар i -д зөв хариулах боломжийг тооцсон 3 параметртай загварыг доор харуулав:

$$T_i(u_i = 1/\theta) = c + \frac{1-c}{1+\exp[-1.7a(\theta-b)]} \quad (1)$$

Таах параметр нь C нь шалгуулагч хэдийгээр даалгаврын хариултыг мэдэхгүй байсан ч зөв хариулах боломжийг харуулж байна. Таах параметр $C=0$ үед хоёр болон гурван хэмжээст загварууд эквивалент байна.

Харин b параметр нь даалгаврын хүнд хөнгөний түвшинг илтгэх бөгөөд шалгуулагчийн зөв хариулах боломж $(.5 + .5c) \cdot (100) \%$ байна.

Даалгаврын параметрууд

Даалгаврын параметр нь IRT-ийн суурь ойлголт юм. Нийтлэг IRT загварууд дээр дурдсанчлан нэг, хоёр, гурван параметртай байдаг. Эдгээр 3 параметрийг дэлгэрэнгүй тайлбарлая.

Даалгаврын дискреминаци буюу “ a ” параметр:

Энэхүү параметр нь өндөр чадвартай шалгуулагчийг сул чадвартай шалгуулагчаас ялгадаг параметр юм. Өөрөөр хэлбэл, даалгавар сайн болон муу шалгуулагчийг хэр сайн ялгаж чадаж

байгааг харуулдаг гэсэн үг юм. Хэрэв “ a ” параметр өндөр байвал өндөр чадвартай шалгуулагч зөв хариулж, бага чадвартай шалгуулагч буруу хариулна гэж ойлгож болно. Даалгаврын дискреминаци -1 ба 1 хооронд байх бөгөөд сайн даалгавруудын “ a ” параметр ихэнхдээ 0.5-1 хооронд байна. Ерөнхийдөө “ a ” параметр 0.4-өөс багагүй байх хэрэгтэй.

Даалгаврын хүнд хөнгөний түвшин буюу “ b ” параметр:

Хэрэв b параметр өндөр байвал тухайн даалгаварт зөв хариулахад хэцүү байна гэсэн үг. Өөрөөр хэлбэл, b параметр өндөртэй даалгаварт чадвар султай шалгуулагчийн зөв хариулах боломж багасна. Харин b параметр бага тохиолдолд ихэнх шалгуулагч, тэр дундаа чадвар султай шалгуулагч ч зөв хариулах боломж байдаг. АНУ-ын Вашингтоны их сургуулиас боловсруулсан ScorePak программын тооцооллоор даалгаврын хүнд хөнгөний түвшин 0.85 болон түүнээс дээш байвал хялбар, 0.51-0.84 хооронд байвал хэвийн, 0.5-аас доош бол хүнд гэж ангилж үздэг.

Таамаглах “ c ” параметр:

Дээр өгүүлсэнчлэн зарим IRT загварууд таамаглах параметртай байдаг. Энэхүү таамаглах c параметр нь сул чадвартай шалгуулагч ч тааж, зөв хариулах боломжтой нөхцлийг тооцоолсон байдаг. Жишээлбэл; 4 хариултын сонголттой даалгаварт шалгуулагчийн тааж, зөв хариулах магадлал 0.25 байна гэсэн үг.

Мэдээлэл цуглуулалт ба түүврийн хэмжээ

Судалгаанд хот, хөдөөгийн ерөнхий боловсролын тус бүр 5 сургуулийн 7 дугаар ангийн 482 сурагчид хамрагдав. Математик чадварыг хэмжээсжүүлэх үүднээс тус бүр 4 сонголт бүхий 30 даалгавартай тест боловсруулсан ба 15 даалгаврыг Олон улсын математик, шинжлэх ухааны хичээлийн сорилоос (TIMSS), 15 даалгаврыг ЕБС-ын

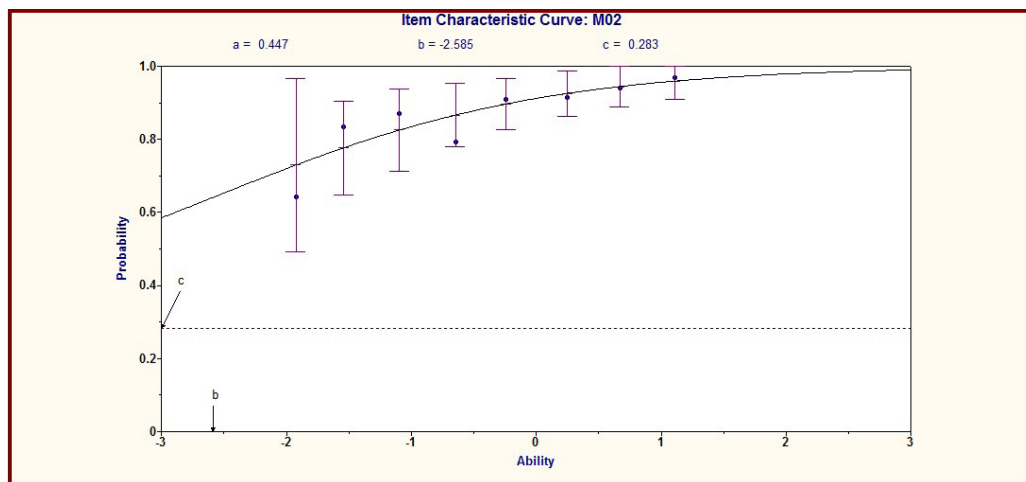
7 дугаар ангийн сурагчдын математикийн хичээлээр эзэмших мэдлэг чадварын хүрээнд боловсруулсан болно. Тестийн даалгавруудыг арифметик, статистик ба магадлал, геометр, алгебр, пропорц гэсэн 5 мэдлэгийн хүрээнд сонгосон бөгөөд даалгаврын тоог дунд сургуулийн хөтөлбөрт заасан математикийн сэдвийн цагаас хамааруулан боловсруулсан. Сурагчид тестийг 40 минутын хугацаанд гүйцэтгэсэн болно. Даалгаврын параметрууд болон даалгаврын онцлогийн муруйг IRT-ийн гол судалгааны хэрэгсэл болох BILOG-MG программын тусламжтайгаар гаргаж авсан болно.

Даалгаварт шинжилгээ хийх

Хөтөлбөрөөр олгож буй мэдлэгийг сурагчид хэрхэн эзэмшиж байна гэдгийг үнэлэхийн тулд тэдгээрийг хэмжихэд чиглэсэн даалгаврууд бүхий тест боловсруулах шаардлага гардаг. Харин

тухайн тестийн даалгаврууд хэр чанартай болсон бэ гэдгийг тогтоохын тулд даалгавруудад анализ хийх нь чухал ач холбогдолтой юм. Тестийн даалгавруудад анализ хийснээр чанартай тест боловсруулж чадна. Тухайлбал, тухайн судлагдахууны хүрээнд анхан шатны мэдлэгийг шалгах даалгаврууд (хэтэрхий хялбар даалгаврууд) олон байвал сурагчид хөөсөрсөн өндөр оноо авах бөгөөд сурагчдын мэдлэгийг хангалттай зөв үнэлж чадахгүй. Харин хэтэрхий хүнд даалгаврууд олон байвал сурагчид маш бага оноо авч улмаар сэдэл, тэмүүлэл нь буурч, суралцах сонирхолгүй болох сөрөг талтай.

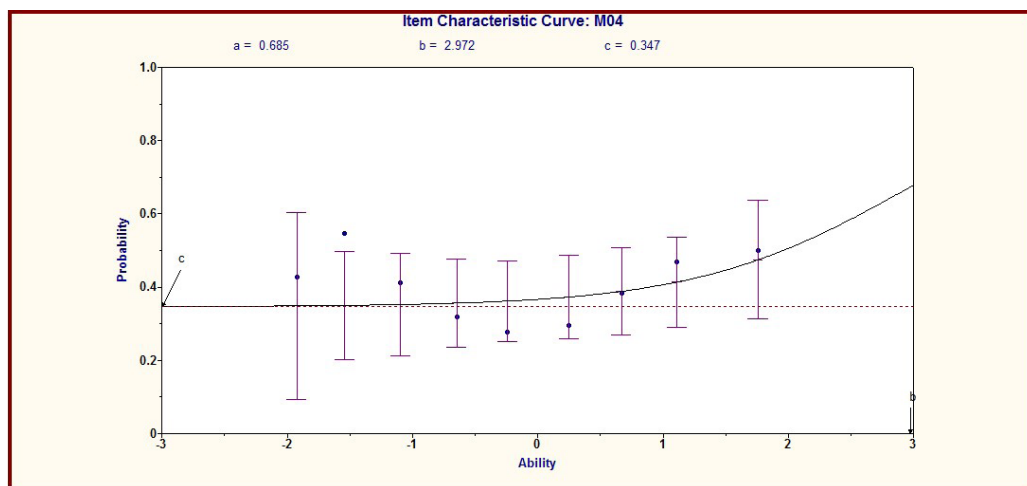
Дээр дурдсан математикийн тестийн даалгавруудад IRT-ийн a , b , c параметрууд болон даалгаврын онцлогийн муруйн (ICC) тусламжтайгаар шинжилгээ хийж үзье. Зураг 1-д даалгавар 2-ын онцлогийн муруйг харуулав.



Зураг 1. Даалгавар 2-ын онцлогийн муруй

Нэлээд хүнд даалгаврын онцлогийн муруй нь ийм хэлбэртэй гардаг. Таалтын буюу “ c ” параметр тодорхой хэмжээгээр байгаа ч даалгаврын хүнд хөнгөний түвшин буюу “ b ” параметр -2,5 байгаа нь энэ даалгавар хэтэрхий хүнд байсныг илтгэж байна. Даалгаврын дискреминаци буюу “ a ” параметр 0,4 байгаа нь сурагчдыг чадварын

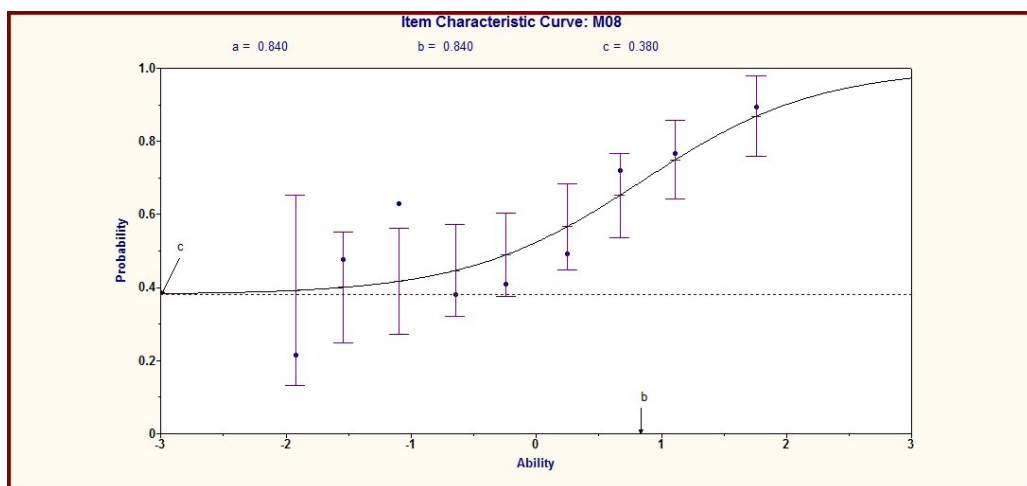
түвшингээр нь сайн ялгаж чадаагүйг илтгэж байна. Дүгнэж хэлбэл энэ даалгавар сайн даалгавар болж чадаагүй бөгөөд дараачийн удаа тест боловсруулахдаа анхаарах, яагаад сурагчид энэ даалгаврыг муу гүйцэтгэв гэдэгт дүгнэлт өгөх шаардлагатайг харуулж байна.



Зураг 2. Даалгавар 4-ийн онцлогийн муруй.

Харин маш хялбар даалгаврын муруй ийм хэлбэртэй байдаг. Таалтын буюу “с” параметр харьцангуй өндөр бөгөөд даалгаврын хүнд хөнгөний түвшин буюу “b” параметр 2,9 байгаа нь энэ даалгавар хэтэрхий хялбар байсныг илтгэж байна. Харин даалгаврын дискреминаци буюу

“a” параметр 0,685 байгаа нь сурагчдыг чадварын түвшингээр нь ялгаж чаджээ. Дүгнэж хэлбэл энэ даалгавар мөн л сайн даалгавар болж чадаагүй бөгөөд үүнээс арай дээр түвшний даалгавар боловсруулах шаардлага байгааг харуулж байна.



Зураг 3. Даалгавар 8-ын онцлогийн муруй.

Боломжийн сайн даалгаврын онцлогийн муруй ихэвчлэн ийм хэлбэртэй харагддаг. Даалгаврын хүнд хөнгөний түвшин буюу “b” параметр 0,84 байгаа нь ScorePak программын тооцооллоор боломжийн гэж үзэж болох ба даалгаврын дискреминаци

буюу “a” параметр 0,84 байгаа нь сурагчдыг чадварын түвшингээр нь сайн ангилж чадсан байна. Ерөнхийдөө энэ даалгавар сайн даалгавар болж чадсан гэж дүгнэж болохоор байна.

Судалгааны ажилд ашигласан тест



хэр найдвартай болсныг мөн шалгаж үзэх хэрэгтэй. Хүснэгт 2-т тестийн даалгавар тус бүрийн найдвартай чанарыг, Хүснэгт

3-т тестийн найдвартай чанарыг тус тус харуулав.

Хүснэгт 2. Тестийн даалгавруудын найдвартай чанарын үзүүлэлт

Даалгавар	Найдвартай чанар	Даалгавар	Найдвартай чанар
1	0.777	16	0.775
2	0.779	17	0.784
3	0.773	18	0.774
4	0.789	19	0.768
5	0.780	20	0.772
6	0.782	21	0.768
7	0.772	22	0.780
8	0.775	23	0.786
9	0.767	24	0.768
10	0.769	25	0.767
11	0.768	26	0.768
12	0.771	27	0.765
13	0.774	28	0.770
14	0.773	29	0.774
15	0.765	30	0.771

Хүснэгт 3. Тестийн найдвартай чанарын үзүүлэлт

Тестийн мэдээлэл	Даалгаврын тоо	Сурагдчын тоо	Найдвартай чанар
Математик	30	482	0.779

Тестийн онолоор тестийн найдвартай чанарын үзүүлэлт 0-1 хооронд байдаг бөгөөд 0.7-оос дээш байх шаардлагатай гэж үздэг. Тестийн найдвартай чанарын үзүүлэлт 0.7- 0.8 байвал сайн тест гэж үзэх бөгөөд магадгүй зарим даалгавруудыг сайжруулах шаардлагатай, 0.8-0.9 байвал

сайн тест, 0.9 буюу түүнээс дээш байвал маш сайн тест гэж үзнэ.(2) Судалгаанд ашигласан тестийн найдвартай чанарын үзүүлэлт 0.7-оос дээш байгаа нь тухайн тестийг хангалттай найдвартай гэж үзэж болохоор байна.

ДҮГНЭЛТ

Боловсролын салбарт үндэсний түвшинд, сургуулийн түвшинд, ангийн түвшинд гэхмэтчилэн сургалтын үйл ажиллагаанд тогтмол үнэлгээ хийдэг байх нь их ач холбогдолтой. Тэдгээр үнэлгээний үр дүнгээс шалтгаалан сургалтын үйл ажиллагаатай холбоотой

чухал шийдвэрүүд гардаг байх ёстой юм. Даалгаврын хариултын онолыг боловсролын салбар тэр дундаа сургалтын үйл ажиллагаанд үнэлгээ хийх томоохон аргазүй гэж ойлгож болно. Манай улсын хувьд элсэлтийн ерөнхий шалгалт болон бусад түвшний тестүүдийн даалгаварт



Даалгаврын хариултын онолыг ашиглаж анализ хийснээр хөтөлбөрөөр олгохыг зорьж буй тодорхой мэдлэг, чадварыг зөв үнэлэх, цаашлаад хөтөлбөрийн агуулгыг сайжруулах арга замыг тодорхойлох, дараа дараачийн тестийн даалгавруудыг сайжруулах боломжтой болох юм. Мөн түүнчлэн багш нар энэ онолыг ашиглан

тест боловсруулах, тестийн даалгавруудад анализ хийх аргазүйд суралцсанаар аль мэдлэгийг хэрхэн заах шаардлагатай байна вэ гэдгийг олж мэдэх, өөрийн хичээл зааж буй арга барилдаа дүн шинжилгээ хийх, сурагчдын мэдлэг, чадварын түвшинг зөв тогтоох боломжтой юм.

Ашигласан бүтээлийн жагсаалт

1. *Deborah Harris. (1989). Comparison of 1-, 2-, and 3- parameter IRT models, Educational Measurement: Practice and Issues, 36-37.*
2. *Nunnally, J. C. (1967). Psychometric theory, 172-235, 6-26, 196. New York: McGraw-Hill.*



SIGNIFICANCE OF USING ITEM RESPONSE THEORY

Abstract

This paper is aimed to explain concept of Item Response Theory (IRT) and introduce effectiveness of this test developing approach. IRT is significantly effective approach for developing and evaluating tests. Therefore IRT can be an appropriate tool to measure examinee ability, test item selection, and test form equating. More significantly, IRT allows teachers and researchers to evaluate examinee's ability and to describe how well items on the test are performing. In addition, IRT uses the concept of Item Characteristic Curve (ICC) to show the relationship between examinee ability and performance on an item. Another important concept of IRT is item parameters. The number of item parameters depends on which IRT model will be used. In this paper 3 parameters logistic model (3PL) was used to estimate students' ability. Research participants were 482 students from secondary schools in Mongolia.